LB 1576
.H4

Class LB1576
Book .H4

PRESENTED BY

A SCALE FOR THE MEASUREMENT OF QUALITY IN ENGLISH COMPOSITION BY YOUNG PEOPLE

613
30

By MILO B. HILLEGAS

Assistant Professor of Elementary Education, Teachers College
Columbia University

Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy, in the Faculty of Philosophy, Columbia University

[Reprinted from TEACHERS COLLEGE RECORD, Vol. 13, No. 4, September, 1912, pp. 331-384.]

PUBLISHED BY
Teachers College, Columbia University
NEW YORK CITY
1912

LB1576
.H4

...versity
OCT 23 19..

ACKNOWLEDGMENTS

At the completion of the study here presented, I find myself deeply obligated to many individuals. The task of judging the sample compositions was not always an agreeable one, and the time necessary to do this work was no small matter with busy people; yet five hundred and fifteen individuals were willing to assist by grading one or the other of the sets of compositions, and thereby made possible the derivation of the scale. The coöperation of many individuals was due to the generous interest manifested by the editors of *Science* and *The Journal of Educational Psychology* in bringing the study to the attention of their readers.

I acknowledge my greatest obligation and deepest gratitude to Dr. E. L. Thorndike with whose advice and encouraging assistance this work was undertaken and conducted.

M. B. H.

CONTENTS

Section		Page
1.	Introduction. (A brief summary of the work that has been done on standards)	1
2.	The Scale. (The scale printed)	5
3.	What the Scale Measures	9
4.	The Location of the Zero Point for English Composition	12
5.	Theory of the Method Employed	14
6.	Sample Compositions from which the Scale was Derived	22
7.	Results Obtained with the First Set of Samples and the Selection of the Second Set	23
8.	Results Obtained with the Second Set of Samples	28
9.	Sufficiency of Judges	50
10.	The Quality of Judges	51
11.	The Use of Mixed Types of Compositions	55

TEACHERS COLLEGE RECORD

Vol. XIII SEPTEMBER, 1912 No. 4

SCALE FOR THE MEASUREMENT OF QUALITY IN ENGLISH COMPOSITION BY YOUNG PEOPLE

Section 1. Introduction

Every attempt to measure the efficiency of instruction in a school system or to evaluate different methods of educational procedure serves to emphasize the importance of standards. Proper standards would make it possible to compare with certainty the work done in one school or system of schools with that done elsewhere. They would make it more difficult for mere opinion to control so much of our school-room practice. After examining the arithmetical abilities of many pupils in different representative school systems, Dr. C. W. Stone ('08) says: "Probably the truest single expression of the findings of this study is summed up in the word *diversity*. . . . Freedom and initiative are here seen to have led educational practice in widely varying paths. Certain paths are those of legitimate differentiation, but others are *waste*. . . . The greatest need shown by this research is standard of achievement. That the great variability herein shown would exist if school authorities possessed adequate means of measuring products is inconceivable, ..."[1]

Standards are greatly needed in order that we may define educational requirements. Results that are accepted in one col-

[1] Arithmetical Abilities and Some Factors Determining Them, Columbia University Contributions to Education, Teachers College Series, No. 19.

lege or school as worthy of a mark of "good," "seventy-five," or any other of the common conventional ratings, may or may not receive the same credit in another educational system. Such differences in the acceptance of results are seldom warranted by the variations in the statements of requirements issued by the institutions concerned. Indeed, as wide divergence in the acceptance of results often exists between schools of the same system, and not infrequently it is found between different departments in the same institution.

If there were standards or scales for the measurements of results in the various school subjects that would approximate the accuracy of the scales used in measuring extension, weight and time, educational administrators and investigators would be able to measure and express the efficiency of a school system in terms that would carry conviction. Such standards would also make it possible to define exactly the requirements of the college, high school, and civil service: and by them school superintendents could define requirements for promotions from one grade to the next.

The more or less satisfactory standards or scales that have been established in certain school subjects show that the task, while difficult, is by no means impossible. Dr. J. M. Rice ('97)[2] tested the spelling ability of some 3,000 public school pupils in 21 different school systems. His purpose was to compare the conditions existing in the various schools and to determine which influenced the spelling ability of the pupils. In order to do this he developed a series of tests which were given in each of the schools under as nearly the same conditions as possible. When the returns from these tests were scored, the tabulated results furnished a standard with which any teacher might compare the work of a class in spelling.

Dr. O. P. Cornman ('02)[3] made a more intensive study of the same character in several of the Philadelphia schools. He used some of the Rice tests and also developed others. The tests

[2] The Futility of the Spelling Grind, *The Forum,* Vol. 23, pp. 163-172 and 409-419.

[3] Spelling in the Elementary School; An Experimental and Statistical Investigation.

were uniformly scored and the results were tabulated in proper form for making comparisons between the various schools.

Dr. Rice ('02)[4] also tested the arithmetical abilities of about 6,000 school children in 17 schools in 7 cities. The methods employed in this research were similar to those employed in the spelling investigation.

Dr. Stone ('08)[5] tested the arithmetical abilities in fundamentals and reasoning of some 6,000 school children in the 6A (high 6th) grade in 26 representative public school systems. Two of the problems on which this study has specific bearing are: "(1) What is the nature of the product of the first six years of arithmetic work? (2) What is the relation between distinctive procedures in arithmetic work and the resulting abilities?" A carefully differentiated test in the fundamentals was used and a separate test was given for reasoning abilities. As was done in all the other investigations, the results were uniformly scored, only in this case greater effort was made to evaluate the different problems. The results are to be found in tabular form. Dr. Stone says of his work, "It is believed that the present study will help to standardize the work in arithmetic in the first six grades. Anyone who wishes may know how his system or school compares with the representative systems of the country." The author gives explicit directions for the use of the tests, the method of scoring the results, and the use of the tables in comparing results. Subsequently these tests have been tried in a number of places, and the results have served to indicate the relative efficiency of the different schools.

The whole process of developing such standards is well shown in the work which Mr. S. A. Courtis ('11)[6] is conducting. He has developed an improved set of arithmetic tests by which it is

[4] Educational Research: A Test in Arithmetic, *The Forum,* Vol. 34, pp. 281-297. Causes of Success and Failure in Arithmetic, *The Forum,* Vol. 34, pp. 437-452.

[5] Arithmetical Abilities and Some Factors Determining Them, Columbia University Contributions to Education, Teachers College Series, No. 19.

[6] Measurements of Growth and Efficiency in Arithmetic, *Elementary School Teacher,* Vol. 10, pp. 58-74 and 177-199. *Elementary School Teacher,* Vol. 12, pp. 127-137. *Journal Educational Psychology,* Vol. 2, p. 272. The Courtis Standard Tests in Arithmetic, with Manual of Instructions for giving and scoring.

claimed any grade in the school may be tested. These have been given to many thousands of children under uniform conditions. The results, when properly tabulated, will form a scale by which it will be possible to measure the achievement of any class or pupil.

The feature of the studies which makes their consideration relevant to this study is not the conclusions reached but the fact that in each case tests were given to a sufficient number of pupils to enable the authors to establish a standard or scale with which they were able to measure the relative efficiency of the various systems.

Dr. E. L. Thorndike ('10)[7] has published a scale for the measurement of quality in the handwriting of children and also one for the handwriting of women. The derivation of these scales involved a somewhat different procedure from that employed for the scales in spelling and arithmetic. Many samples both of children's handwriting and of adults' were secured and these were judged separately by many competent individuals. From the scores thus obtained the relative values of the samples were determined and the scales formed. The scales in this case consist of a series of samples of handwriting with which the samples that are to be measured are compared. In many particulars the methods employed in deriving these scales in handwriting are the same as those used in the present study, and frequent reference will be made to Dr. Thorndike's study in the descriptions which follow.

[7] Handwriting, *Teachers College Record,* Vol. 11, No. 2.

Section 2. The Scale

This scale for the measurement of quality in English composition by young people has been derived by methods that are described in the following pages.

Sample 580. **Value 0.** Artificial sample.

Letter.

Dear Sir: I write to say that it aint a square deal Schools is I say they is I went to a school. red and gree green and brown aint it hito bit I say he don't know his business not today nor yeaterday and you know it and I want Jennie to get me out.

Sample 595. **Value 183.** Artificial sample.

My Favorite Book.

the book I refer to read is Ichabod Crane, it is an grate book and I like to rede it. Ichabod Crame was a man and a man wrote a book and it is called Ichabod Crane i like it because the man called it ichabod crane when I read it for it is such a great book.

Sample 618. **Value 260.** Artificial sample.

The Advantage of Tyranny.

Advantage evils are things of tyranny and there are many advantage evils. One thing is that when they opress the people they suffer awful I think it is a terrible thing when they say that you can be hanged down or trodden down without mercy and the tyranny does what they want there was tyrans in the revolutionary war and so they throwed off the yok.

Sample 94. **Value 369.** Written by a boy in the second year of the high school, aged 14 years.

Sulla as a Tyrant.

When Sulla came back from his conquest Marius had put himself consul so sulla with the army he had with him in his conquest siezed the government from Marius and put himself in consul and had a list of his enemys printy and the men whoes names were on this list we beheaded.

Sample 519. **Value 474.** Written by a girl in the third year of the high school, aged 17 years.

De Quincy.

First: De Quincys mother was a beautiful women and through her De Quincy inhereted much of his genius.

His running away from school enfluenced him much as he roamed through the woods, valleys and his mind became very meditative.

The greatest enfluence of De Quincy's life was the opium habit. If it was not for this habit it is doubtful whether we would now be reading his writings.

His companions during his college course and even before that time were great enfluences. The surroundings of De Quincy were enfluences. Not only De Quincy's habit of opium but other habits which were peculiar to his life.

His marriage to the woman which he did not especially care for.

The many well educated and noteworthy friends of De Quincy.

Sample 534. **Value 585.** Written by a boy in the fourth year of the high school, aged 16 years.

Fluellen.

The passages given show the following characteristic of Fluellen: his inclination to brag, his professed knowledge of History, his complaining character, his great patriotism, pride of his leader, admired honesty, revengeful, love of fun and punishment of those who deserve it.

Sample 196. **Value 675.** Written by a girl in the first year of the high school, aged 18 years.

Ichabod Crane.

Ichabod Crane was a schoolmaster in a place called Sleepy Hollow. He was tall and slim with broad shoulders, long arms that dangled far below his coat sleeves. His feet looked as if they might easily have been used for shovels. His nose was long and his entire frame was most loosely hung to-gether.

Sample 221. **Value 772.** Written by a boy in the third year of the high school, aged 16 years.

Going Down with Victory.

As we road down Lombard Street, we saw flags waving from nearly every window. I surely felt proud that day to be the driver of the gaily decorated coach. Again and again we were cheered as we drove slowly to the postmasters, to await the coming of his majestie's mail. There wasn't one of the gaily bedecked coaches that could have compared with ours, in my estimation. So with waving flags and fluttering hearts we waited for the coming of the mail and the expected tidings of victory.

When at last it did arrive the postmaster began to quickly sort the bundles, we waited anxiously. Immediately upon receiving our bundles, I lashed the horses and they responded with a jump. Out into the country we drove at reckless speed—everywhere spreading like wildfire the news, "Victory!" The exileration that we all felt was shared with the horses. Up and down grade and over bridges, we drove at breakneck speed and spreading the news at every hamlet with that one cry "Victory!" When at last we were back home again, it was with the hope that we should have another ride some day with "Victory."

Sample 571. **Value 838.** Written by a boy in the Freshman class in college.

Venus of Melos.

In looking at this statue we think, not of wisdom, or power, or force, but just of beauty. She stands resting the weight of her body on one foot, and advancing the other (left) with knee bent. The posture causes the figure to sway slightly to one side, describing a fine curved line. The lower limbs are draped but the upper part of the body is uncovered. (The unfortunate loss of the statue's arms prevents a positive knowledge of its original attitude.) The eyes are partly closed, having something of a dreamy langour. The nose is perfectly cut, the mouth and chin are moulded in adorable curves. Yet to say that every feature is of faultless perfection is but cold praise. No analysis can convey the sense of her peerless beauty.

Sample 177. **Value 937.** Written by a boy in the Freshman class in college.

A Foreigner's Tribute to Joan of Arc.

Joan of Arc, worn out by the suffering that was thrust upon her, nethertheless appeared with a brave mien before the Bishop of Beauvais. She knew, had always known that she must die when her mission was fulfilled and death held no terrors for her. To all the bishop's questions she answered firmly and without hesitation. The bishop failed to confuse her and at last condemned her to death for heresy, bidding her recant if she would live. She refused and was lead to prison, from there to death.

While the flames were writhing around her she bade the old bishop who stood by her to move away or he would be injured. Her last thought was of others and De Quincy says, that recant was no more in her mind than on her lips. She died as she lived, with a prayer on her lips and listening to the voices that had whispered to her so often.

The heroism of Joan of Arc was wonderful. We do not know what form her great patriotism took or how far it really led her. She spoke of hearing voices and of seeing visions. We only know that she resolved to save her country, knowing though she did so, it would cost her her life. Yet she never hesitated. She was uneducated save for the lessons taught her by nature. Yet she led armies and crowned the dauphin, king of France. She was only a girl, yet she could silence a great bishop by words that came from her heart and from her faith. She was only a woman, yet she could die as bravely as any martyr who had gone before.

Section 3. What the Scale Measures

The scale is composed of sample English compositions, the qualities of which have been determined by more than four hundred competent judges. The values which are assigned to the various samples express their quality in the same sense, though not as accurately, as millimeters express the lengths of lines. Just as 183 mm. may be expressed by 18.3 cm. or 1.83 dm., so the 183 employed in this scale may be considered as 183 small units of quality, or as 18.3 units ten times as large, or as 1.83 units one hundred times as large. The values here employed should not be confused with the customary per cents that are used in rating English compositions. A difference of two hundred in this scale is equal to twice a difference of one hundred taken in any part of the scale. Thus sample 94 with a value of 369 is a little more than twice as good an English composition as sample 595 with a value of 183. Sample 519 with a value of 474 is about one-half as good as sample 177 with a value of 937.

Merit in English writing is complex. Judges are influenced both by form and by content. Such factors of form as spelling, punctuation, capitalization, and the like are subject to definite rules. Form is, therefore, more easily measured than content. When an individual is in doubt concerning the relative merits of two English compositions, the tendency is to fix upon one or more of the obvious form elements and for the time being to give them undue importance in fixing the relation of the samples.

No attempt has been made in this study to define merit. The term as here used means just that quality which competent persons commonly consider as merit, and the scale measures just this quality. The accompanying graphic representation of the location of the various samples composing the scale will serve to make plain the meaning of the scale.

Any specimen of English composition may be measured by the scale by placing the sample to be measured alongside the samples constituting the scale and determining to which it most nearly corresponds. The use of the scale will be clearly under-

stood, if the reader will compare sample 376, given below, with the scale. There will be no hesitation in pronouncing sample 376 better than the lower end of the scale until quality 585 is reached. The majority of persons will consider sample 376

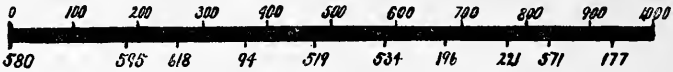

FIG. I. Graphic representation of the relative values of the samples composing the scale. The numbers above the line represent the various unit points in the scale. The numbers below the line represent sample compositions and show their relative positions on the scale. Thus, sample 580 is at the zero point, sample 595 is between the values of 100 and 200, etc.

worth about 772 or the equal of sample 221. In an actual test with seventy-three individuals, seventeen considered sample 376 inferior to sample 534, sixteen thought it inferior to sample 196, and twenty-one thought sample 376 was better than sample 571; but there was almost perfect agreement that sample 376 was nearly if not quite equal to sample 221. If the reader thinks sample 376 is better than sample 221 but not as good as sample 571, he may place the value between 772 and 838. By this method the value of any sample may be expressed as accurately as the individual cares to make it. "The sample to be measured should, for convenience, be examined with the entire scale in view. If the scale's samples are arranged in order on a table or against a wall, the examined sample is easily compared with them. The measurer then decides what quality of the scale the sample possesses and records the measure. . . . The measure may be made more and more accurate by having other judges also measure, each always in ignorance of the ratings given by the others. In default of other judges, the measure may be made more accurate by rating the sample two or more times, each time in ignorance of the ratings previously given. An individual may be measured more accurately by using several samples of his compositions, each being rated in ignorance of the ratings given to the other samples."[8]

[8] Handwriting, p. 8.

Sample 376. The Military Career of Alexander the Great.

Alexander the Great one of the greatest generals of the world began his military career at the age of twenty. It was then at his father's death that he crushed all rebellion both in Macedonia and Greece. Then he started on his famous march. Crossing the Hellespont he defeated the Persians at the river Granicus from thence he followed the coast of Asia Minor. However he turned aside from his course and going up into Phrygium he cut the Gordian knot. Next he defeated a large Persian army at Issus, captured Tyre after a long siege and much labor. He now went into Egypt, founded Alexandria, and returned to Persia. At Arbela the Persian army was annihilated and Darius fled. Alexander next conquered Babylon and Persepolis and then having wandered thru eastern Persia he crossed the Hindu Hush Mountains and entered India. Although victorious Alexander's army was weary of such hard travelling and refused to go farther. So the great general was forced to lead them back to Babylon and his military career was finished.

Section 4. The Location of the Zero Point in English Composition[9]

The location of the zero point for merit in English composition by young people is of importance in order to allow the "*times as much merit*" judgment. Measures of fatigue, practice, or change of any sort will be greatly facilitated if we can so arrange the scale that 8, 10, and 12 on it mean twice as far from just not any of the thing in question as 4, 5, and 6; that 240 means twice as much of the thing as 120, three times as much as 80, five times as much as 48. Hitherto no one could say with any assurance whether sample 571 was two, ten or a hundred times as "good" as 618.

It has been located approximately as sample 580 on the basis of the judgments of (1) nine men of special literary ability, five of them professors of English, of whom four are also authors of standard text-books on English composition, and four men of marked general intellect and literary ability, (2) eleven gifted teachers familiar with secondary education, and (3) eight psychologists familiar with the significance of scales and zero points in the case of intellectual abilities and products.

Although no one person of these had any deliberate criterion for the point where positive merit just begins, and although one's first reaction to the request to locate such a point is to regard it as arbitrary, there is much agreement among individuals and almost perfect agreement in the case of the averages of the three groups.

These individuals located the zero point in connection with the set of 27 samples[10] either by placing it below the worst of these or by stating which of these they considered to possess absolutely no merit as specimens of English composition by young people.

Of the twenty-eight judges, two regard zero merit as something below the least meritorious specimens of this list, and five put it higher than sample 618, but three-fourths of the

[9] For the experimentation and calculation in connection with the location of the zero point, Professor E. L. Thorndike is responsible.

[10] For the samples used in this work see Section 8.

judges locate it as lower than sample 607 or higher than sample 595. The central tendency of the twenty-eight puts it at or just barely below specimen 580.

Taking the sample regarded as just barely of some merit and that regarded as the next below it in the case of each judge, we have the following results:

Below 607	4 times	Below 580	5 times
" 607	9 "	" 39	3 "
" 670	2 "	" 79	6 "
" 58	3 "	" 595	2 "
" 501	6 "	" 603	2 "
" 491	3 "	" 618	2 "
		Above 618	9 "

The median opinion of these 28 judges thus places the beginning of merit in English writing by young people at sample 580.

Section 5. Theory of the Method Employed in Deriving the Scale

Nearly five hundred individuals judged one or the other of the sets which contained the samples used in the scale. Each person was asked to arrange the samples in order of merit. This amounted to a request that each sample should be compared with every other sample in the set. In scoring the results, the poorest sample was numbered one, the next poorest two, and so on; thus in a set of eighty-three samples the best was numbered eighty-three. The scale was derived from these scores by the method of right and wrong cases.[11] The theory of this method as applied to this study may be stated as follows: Differences that are equally often noticed are equal, unless the differences are either always or never noticed.[12]

Merit in English writing is the resultant of a large number of independent factors. Among the individuals who judged the samples used in this study, several greatly undervalued certain samples because they thought that the language used was too mature for high school pupils. On the other hand, several of the artificial samples were greatly overvalued because the judges thought that they detected childish modes of expression in these. When a large number of individuals are required to judge the quality of any sample English composition, many will be nearly correct, some will be in error because of overvaluing it, and an equal number will undervalue it. Small errors of judgment will be much more common than large ones. The theory of probability enables us to construct a diagram that will represent what would result if a large number of persons were to judge an English composition, or what would result if one person were to judge independently a large number of times. Figure II shows the probable grouping of judgments, when a large number of individuals are requested to determine the value of a composition. The line XO diagrammatically represents the value of an English composition. The bell-shaped surface, ABC, shows the distribution of judgments. The shape of the

[11] Fullerton and Cattel, On the Perception of Small Differences, pp. 12 ff.

[12] Handwriting, p. 5.

curve will depend upon the accuracy of the judges. The better the judges the higher that part of the curve above O and the shorter the base AC. The surface ABC represents diagrammatically a normal surface of distribution. The line BO divides the surface into two equal parts. PQ and MN are equally distant from BO and within the area PMNQ, fifty per cent of the judgments are found.[13] The line OP or its equal OM represents an error of such size that half of the errors are larger and half are smaller. It is the median error or as it is called the *median deviation* (M. D.)[14]

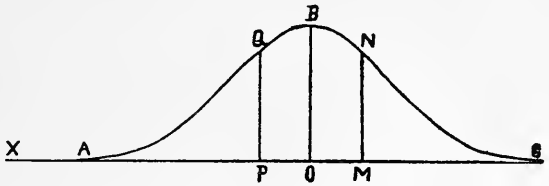

FIG. II. A normal surface of frequency on which is shown the distribution of errors in judgment when a large number of individuals are required to determine the value of a sample English composition whose value by concensus of opinion is represented by the line XO.

In Fig. III, the values of two samples are represented, one by the line XO and the other by the line X′O′. Such a difference in value is taken that the line O′P′, which represents this difference, is equal to the median deviation. The area P′C′Q′ contains seventy-five per cent of the judgments of the sample represented by X′O′. Seventy-five per cent, therefore, of the judgments are correct and twenty-five per cent (the area A′P′Q′) are wrong, because they make the sample represented by XO better than the sample represented by X′O′. When seventy-five per cent of the judges notice the difference in quality between two samples, the difference is equal to the median deviation, and the ratio between the difference and the median deviation in this case is 1/1 or 1. Throughout this study this ratio is taken as the unit of value. The unit may be

[13] For a complete statement of the relations that exist between the various parts of a normal surface of distribution, see Thorndike, Mental and Social Measurements, p. 59.

[14] The name 'probable error' is applied to this value.

defined as that difference in quality which exactly seventy-five per cent of the judges observe.

In Figs. IV and V the arrangement of the curves is such as to show the relation that exists between the difference and the median deviation when various percentages of the judges notice the difference. Curves No. 1 and No. 2 show this relation when

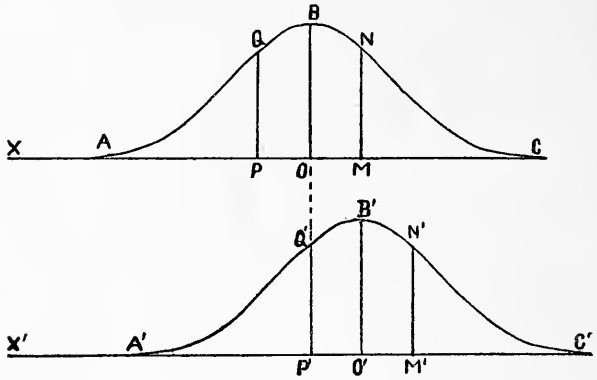

FIG. III. Two surfaces of frequencies so placed as to show the difference in value of two samples of English compositions when seventy-five per cent of the judges consider one better than the other.

fifty per cent of the judges believe that sample represented by curve No. 2 is better than sample represented by curve No. 1. In this case BO extended passes through point O′ and there is no difference in the quality of these samples. Curve No. 3 is so placed that fifty-five per cent of the judges consider sample represented by curve No. 3 better than sample represented by curve No. 1. The difference is represented by the line Y′O″. By actual measurement Y′O″ is $\frac{19}{100}$ of P″O″. Curve No. 4 shows sixty per cent "better" judgments. The difference Y‴O‴ is $\frac{38}{100}$ of the median deviation (P‴O‴). In curve No. 5 the percentage of "better" judgments is sixty-five and the difference is $\frac{57}{100}$ of the median deviation. The following table shows the curve compared with curve No. 1, the percentage of better judgments represented by the curve, and the ratio between the difference in value and the median deviation $\left(\frac{D.}{M.\ D.}\right)$. Median deviation is defined on page 15.

TABLE I

THE RATIOS BETWEEN THE DIFFERENCES IN VALUE AND THE MEDIAN DEVIATION CORRESPONDING TO CERTAIN PERCENTAGES OF "BETTER" JUDGMENTS AS SHOWN BY THE CURVES IN FIGS. III, IV AND V

	Curve compared with curve No. 1	Percentage of "better" judgments	Difference divided by the median deviation (D/M. D.)
Fig. IV.....	No. 2	50	0
"	No. 3	55	.19
"	No. 4	60	.38
"	No. 5	65	.57
"	No. 6	70	.78
Fig. III.....	No. 2	75	1.00
Fig. V.....	No. 7	80	1.25
"	No. 8	85	1.53
"	No. 9	90	1.89
"	No. 10	95	2.43
"	No. 11	100	?

A table giving the differences for all percentages from fifty to ninety-nine is found on page 16 of "On the Perception of Small Differences." This table is also found on page 164 of "Mental and Social Measurements." In selecting the samples that should be used in the scale, it was more convenient to begin with the poorer samples, so the table was transmuted to give the differences for percentages from 1 to 50, and the differences for tenths of a per cent were interpolated. The results are found in Table II.

Any standard or scale should be based on an unit such that equal units may be derived independently of the scale. The unit in this scale has been defined as that difference which seventy-five per cent of the judges are able to distinguish. All that is required to derive this unit is a set of samples that vary from each other by small degrees in quality. When two samples are found such that seventy-five per cent of the judges agree in calling one better than the other, the difference is just the difference used as the unit in this scale.

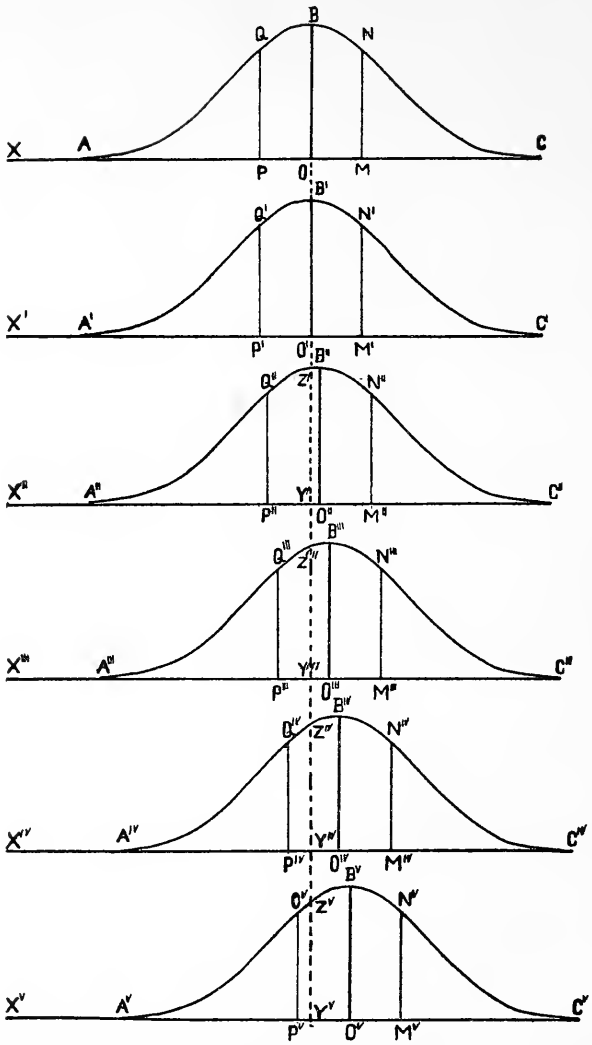

FIG. IV. Various frequency curves so placed in relation to curve number one as to show the difference in value between two samples when the percentage of "better" judgments is either 50, 55, 60, 65, or 70 respectively.

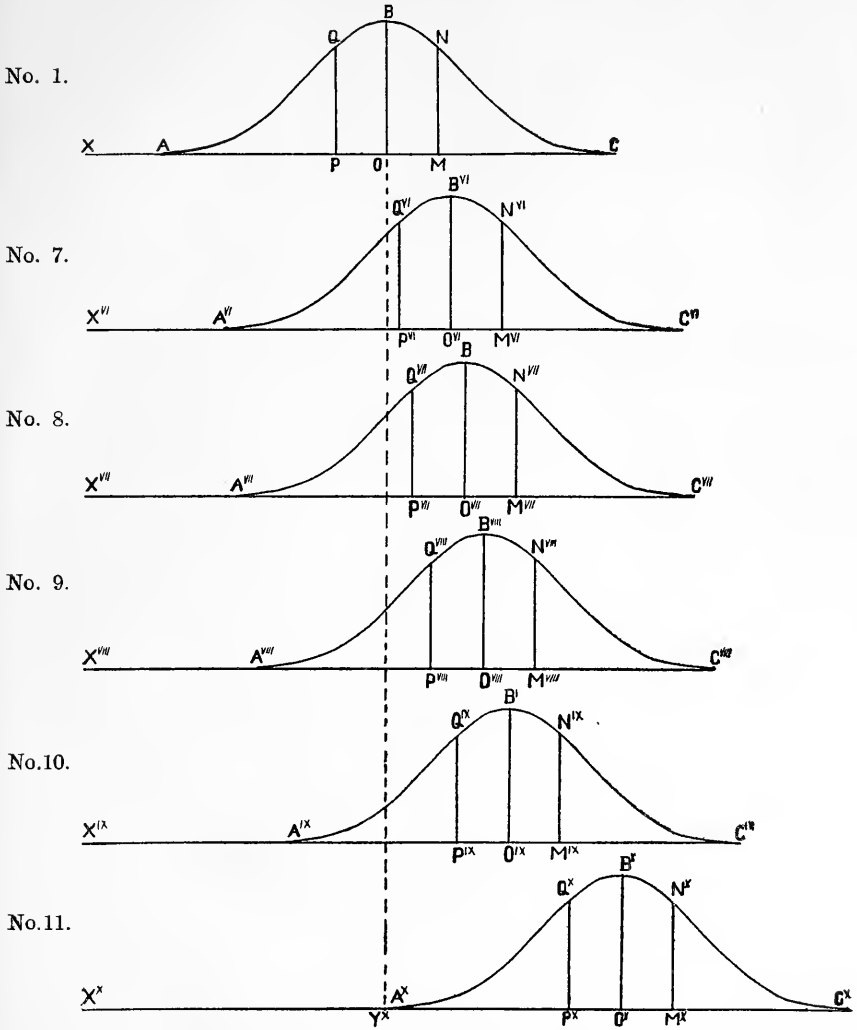

FIG. V. Various frequency curves so placed in relation to curve number one as to show the difference in value between two samples when the percentage of "better" judgment is either 80, 85, 90, 95, or 100 respectively.

TABLE II

TABLE FOR DETERMINING THE MEDIAN DEVIATION FROM THE PERCENTAGE OF WRONG CASES AND THE AMOUNT OF DIFFERENCE

The table reads as follows: When 1 per cent of the judges regard sample A worse than sample B, the difference, A-B, is 3.45 of the median deviation. When 1.1 per cent of the judges regard sample A worse than sample B, the difference, A-B, is 3.41 of the median deviation. When 1.2 per cent of the judges regard sample A worse than sample B, the difference, A-B, is 3.37 of the median deviation, etc.

Percentage of "worse" judgments	D / M. D.	Percentage of "worse" judgments	D / M. D.	Percentage of "worse" judgments	D / M. D.	Percentage of "worse" judgments	D / M. D.	Percentage of "worse" judgments	D / M. D.
1.0	3.45	6.0	2.31	11.0	1.82	16.0	1.47	21.0	1.20
.1	3.41	.1	2.30	.1	1.81	.1	1.46	.1	1.19
.2	3.37	.2	2.29	.2	1.80	.2	1.46	.2	1.19
.3	3.33	.3	2.27	.3	1.80	.3	1.45	.3	1.18
.4	3.29	.4	2.26	.4	1.79	.4	1.45	.4	1.18
.5	3.25	.5	2.25	.5	1.78	.5	1.44	.5	1.17
.6	3.21	.6	2.24	.6	1.77	.6	1.43	.6	1.16
.7	3.17	.7	2.23	.7	1.76	.7	1.43	.7	1.16
.8	3.13	.8	2.21	.8	1.76	.8	1.42	.8	1.15
.9	3.09	.9	2.20	.9	1.75	.9	1.42	.9	1.15
2.0	3.05	7.0	2.19	12.0	1.74	17.0	1.41	22.0	1.14
.1	3.02	.1	2.18	.1	1.73	.1	1.41	.1	1.14
.2	3.00	.2	2.17	.2	1.73	.2	1.40	.2	1.13
.3	2.97	.3	2.16	.3	1.72	.3	1.40	.3	1.13
.4	2.95	.4	2.15	.4	1.71	.4	1.39	.4	1.12
.5	2.92	.5	2.14	.5	1.71	.5	1.39	.5	1.12
.6	2.89	.6	2.12	.6	1.70	.6	1.38	.6	1.12
.7	2.87	.7	2.11	.7	1.69	.7	1.38	.7	1.11
.8	2.84	.8	2.10	.8	1.68	.8	1.37	.8	1.11
.9	2.82	.9	2.09	.9	1.68	.9	1.37	.9	1.10
3.0	2.79	8.0	2.08	13.0	1.67	18.0	1.36	23.0	1.10
.1	2.77	.1	2.07	.1	1.66	.1	1.35	.1	1.10
.2	2.75	.2	2.06	.2	1.66	.2	1.35	.2	1.09
.3	2.73	.3	2.05	.3	1.65	.3	1.34	.3	1.09
.4	2.71	.4	2.04	.4	1.64	.4	1.34	.4	1.08
.5	2.70	.5	2.03	.5	1.64	.5	1.33	.5	1.08
.6	2.68	.6	2.03	.6	1.63	.6	1.32	.6	1.07
.7	2.66	.7	2.02	.7	1.62	.7	1.32	.7	1.07
.8	2.64	.8	2.01	.8	1.61	.8	1.31	.8	1.06
.9	2.62	.9	2.00	.9	1.61	.9	1.31	.9	1.06
4.0	2.60	9.0	1.99	14.0	1.60	19.0	1.30	24.0	1.05
.1	2.58	.1	1.98	.1	1.59	.1	1.30	.1	1.05
.2	2.57	.2	1.97	.2	1.59	.2	1.29	.2	1.04
.3	2.55	.3	1.96	.3	1.58	.3	1.29	.3	1.04
.4	2.54	.4	1.95	.4	1.58	.4	1.28	.4	1.03
.5	2.52	.5	1.95	.5	1.57	.5	1.28	.5	1.03
.6	2.50	.6	1.94	.6	1.56	.6	1.27	.6	1.02
.7	2.49	.7	1.93	.7	1.56	.7	1.27	.7	1.02
.8	2.47	.8	1.92	.8	1.55	.8	1.26	.8	1.01
.9	2.46	.9	1.91	.9	1.55	.9	1.26	.9	1.01
5.0	2.44	10.0	1.90	15.0	1.54	20.0	1.25	25.0	1.00
.1	2.43	.1	1.88	.1	1.53	.1	1.25	.1	1.00
.2	2.41	.2	1.88	.2	1.53	.2	1.24	.2	.99
.3	2.40	.3	1.87	.3	1.52	.3	1.24	.3	.99
.4	2.39	.4	1.86	.4	1.51	.4	1.23	.4	.98
.5	2.38	.5	1.85	.5	1.51	.5	1.23	.5	.98
.6	2.36	.6	1.85	.6	1.50	.6	1.22	.6	.97
.7	2.35	.7	1.84	.7	1.49	.7	1.22	.7	.97
.8	2.34	.8	1.83	.8	1.48	.8	1.21	.8	.96
.9	2.32	.9	1.83	.9	1.48	.9	1.21	.9	.96

Percentage of "worse" judgments	D M. D.	Percentage of "worse" judgments	D M. D.	Percentage of "worse" judgments	D M. D.	Percentage of "worse" judgments	D M. D.	Percentage of "worse" judgments	D M. D.
26.0	.95	31.0	.74	36.0	.53	41.0	.34	46.0	.15
.1	.95	.1	.74	.1	.53	.1	.34	.1	.15
.2	.94	.2	.73	.2	.52	.2	.33	.2	.14
.3	.94	.3	.73	.3	.52	.3	.33	.3	.14
.4	.93	.4	.72	.4	.51	.4	.32	.4	.13
.5	.93	.5	.72	.5	.51	.5	.32	.5	.13
.6	.93	.6	.71	.6	.51	.6	.32	.6	.13
.7	.92	.7	.71	.7	.50	.7	.31	.7	.12
.8	.92	.8	.70	.8	.50	.8	.31	.8	.12
.9	.91	.9	.70	.9	.49	.9	.30	.9	.11
27.0	.91	32.0	.69	37.0	.49	42.0	.30	27.0	.11
.1	.91	.1	.69	.1	.49	.1	.30	.1	.11
.2	.90	.2	.68	.2	.48	.2	.29	.2	.10
.3	.90	.3	.68	.3	.48	.3	.29	.3	.10
.4	.89	.4	.67	.4	.47	.4	.28	.4	.09
.5	.89	.5	.67	.5	.47	.5	.28	.5	.09
.6	.88	.6	.67	.6	.47	.6	.28	.6	.09
.7	.88	.7	.66	.7	.46	.7	.27	.7	.08
.8	.87	.8	.66	.8	.46	.8	.27	.8	.08
.9	.87	.9	.65	.9	.45	.9	.26	.9	.07
28.0	.86	33.0	.65	38.0	.45	43.0	.26	48.0	.07
.1	.86	.1	.65	.1	.45	.1	.26	.1	.07
.2	.85	.2	.64	.2	.44	.2	.25	.2	.06
.3	.85	.3	.64	.3	.44	.3	.25	.3	.06
.4	.84	.4	.63	.4	.43	.4	.24	.4	.06
.5	.84	.5	.63	.5	.43	.5	.24	.5	.06
.6	.84	.6	.63	.6	.43	.6	.24	.6	.05
.7	.83	.7	.62	.7	.42	.7	.23	.7	.05
.8	.83	.8	.62	.8	.42	.8	.23	.8	.05
.9	.82	.9	.61	.8	.41	.9	.22	.9	.04
29.0	.82	34.0	.61	39.0	.41	44.0	.22	49.0	.04
.1	.82	.1	.61	.1	.41	.1	.22	.1	.04
.2	.81	.2	.60	.2	.40	.2	.21	.2	.03
.3	.81	.3	.60	.3	.40	.3	.21	.3	.03
.4	.80	.4	.59	.4	.40	.4	.21	.4	.02
.5	.80	.5	.59	.5	.40	.5	.21	.5	.02
.6	.80	.6	.59	.6	.39	.6	.20	.6	.02
.7	.79	.7	.58	.7	.39	.7	.20	.7	.01
.8	.79	.8	.58	.8	.39	.8	.20	.8	.01
.9	.78	.9	.57	.9	.38	.9	.19	.9	.00
30.0	.78	35.0	.57	40.0	.38	45.0	.19	50.0	.00
.1	.78	.1	.57	.1	.38	.1	.19		
.2	.77	.2	.56	.2	.37	.2	.18		
.3	.77	.3	.56	.3	.37	.3	.18		
.4	.76	.4	.55	.4	.36	.4	.17		
.5	.76	.5	.55	.5	.36	.5	.17		
.6	.76	.6	.55	.6	.36	.6	.17		
.7	.75	.7	.54	.7	.35	.7	.16		
.8	.75	.8	.54	.8	.35	.8	.16		
.9	.74	.9	.53	.9	.34	.9	.15		

Section 6. The Sample Compositions from Which the Scale was Derived

The first step in the derivation of the scale was the collection of about seven thousand English compositions by young people. These were obtained from various sources and represent a definite attempt to obtain particularly the very poorest and the best work that is done in the schools. After collecting these samples the author and one other graded them roughly into ten classes. From these classes seventy-five samples were selected. In order that the samples at the extremes of the scale might be measured, it was necessary to supply some artificial samples. The poorest of these were conscious attempts by adults to write very poor English. The best samples were obtained from youthful writings of such literary geniuses as Jane Austen and the Brontës. The works of some college Freshmen were also used. As finally constituted, the set consisted of eighty-three samples which varied from the poorest to the best by small degrees of quality.

Nearly all of the samples were in the handwriting of the authors, some of which was very good and some very poor. In order that the character of the handwriting might not influence the judges, all the samples were typewritten and mimeographed. Great care was exercised in reproducing all of the mistakes that appeared in the originals.

It would have been desirable to have included more than the eighty-three samples, but the persons whose services were particularly desired were busy, and since the judging of a set of even eighty-three required from three to six hours of close application, it was not possible to enlarge the set.

Separate sets of these samples were given to about one hundred individuals with the following printed request: "Please arrange these in order of merit as specimens of English composition by young people. If there are two or more of them which seem to you to be absolutely equal in merit, give them the same number as a score. Score your results on the accompanying sheet calling the worst specimen 1, the next worst 2, the next worst 3, and so on."

Section 7. The Results Obtained with the First Set of Samples and the Selection of a Second Set

More than one hundred individuals worked with the set of eighty-three samples. Some, however, did not understand the directions and some made such serious mistakes in recording their results that it was possible to use only seventy-three of the records. These records were copied on a large sheet for convenience in handling, and the distribution of the positions assigned to each sample was determined. The next step was to compare the position assigned each sample with that assigned each other sample in order to determine just how many of the judges had considered any given sample better than any other given sample. These "better" judgments were arranged in tabular form such as is shown in Table IV on page 26.

There were great variations in the positions assigned to the samples. It was not possible to establish the position of any sample with reasonable accuracy until many more judgments could be secured. The judgments of the seventy-three individuals indicated the samples that were probably nearly equal in merit, and in this way made it possible to select from the eighty-three samples a much smaller set which would still contain all the important steps in quality from the poorest to the best. The samples that were included in the new set are reproduced in Section 8, and reference to Table V, which shows the percentage of "better" judgments for the samples selected from the set of eighty-three, will help in understanding the method that was employed in selecting this set.

All but three of the judges had agreed that sample 607 was the poorest. This sample was included, as well as the next samples in merit, 580 and 595. The remaining eighteen samples were selected by taking the sample that about seventy-five per cent of the judges had agreed was better than the last sample selected. The differences in merit between samples 607 and 580, and between samples 580 and 595 were large. New artificial samples were included with these three samples and these were judged by a number of individuals in order to determine several samples that would range in merit between the

two pairs. In this way five were selected, making in all a set of twenty-seven samples. Later samples 519 and 520 were added.

Since only twenty-three of the eighty-three samples in the first set are used in any succeeding set, only such data as bear upon these twenty-three samples are here given. Table III shows the distribution of positions assigned to these samples by seventy-four individuals. This table is followed by Table IV which gives the number of "better" judgments for this set, and by Table V which shows the percentage of "better" judgments for each pair of samples. Only the significant data are given in Tables IV and V.

TABLE III

DISTRIBUTION OF THE POSITIONS ASSIGNED TO TWENTY-THREE SAMPLES JUDGED BY SEVENTY-FOUR * INDIVIDUALS

These scores were obtained from the positions assigned these samples when they were in the set of eighty-three. The table should be read as follows: Seventy-two individuals considered sample 607, the poorest; one considered sample 595, the poorest; sixty-five considered sample 580, next to the poorest, etc.

Positions	Samples																						
	607	580	595	618	603	570	94	627	519	534	322	200	196	323	221	300	520	571	220	434	565	488	177
1	72		1	1																			
2	1	65	5	1	1	1																	
3	1	4	40	17	7	3	2																
4		2	22	25	15	8	1	1															
5		2	5	9	33	11	6	6	2														
6		1	1	10	4	20	11	16	5	1			1										
7				5	5	17	17	16	3	5	3		1										
8				4	3	7	21	18	10	4	3	3	1	1									
9				2	1	4	10	9	16	7	9	6	5		1		1		1				
10					1	1	4	4	18	13	10	10	2	2	4	2	2		1			3	
11						2	2	3	8	11	17	9	10	4	5	4		1	2	1			
12								1	4	5	7	21	7	7	2	3	5	1	3	2		1	2
13									4	11	8	5	14	4	10	4	4	2	3	1		1	2
14									2	6	3	6	10	12	4	9	5	5	3	3	2	3	
15									1	1	3	3	7	12	7	6	10	5	6	3	5	3	2
16										5	3	2	5	5	10	11	6	8	7	4	7	1	
17										1	1	5	2	9	10	10	5	8	7	2	6	3	5
18										1	3	2	5	2	7	6	4	15	11	6	2	5	5
19										1		1		8	4	6	11	10	8	9	8	3	5
20										1	3	1	2	2	4	4	7	6	11	9	8	9	6
21									1				1	3	3	4	5	5	4	11	13	10	15
22										1	1		1	1	3	2	4	6	5	13	10	11	16
23														2		3	5	2	2	10	13	21	16

* Not all the individuals judged each sample.

TABLE IV

THE NUMBER OF "BETTER" JUDGMENTS OF SEVENTY-FOUR INDIVIDUALS WITH A SET OF TWENTY-THREE SAMPLES SELECTED FROM THE SET OF EIGHTY-THREE

The table should be read as follows: None of the seventy-four judges considered sample 607 better than sample 580, one judge considered sample 607 better than 595, nine judges considered sample 580 better than sample 595, etc.

Sample Compared	Samples with which comparison is made																					
	580	595	618	603	570	94	627	519	534	322	200	196	323	221	300	520	571	220	434	565	488	177
607	0	1	1	0	1																	
580		9	3	3	3																	
595			19	11	3	1																
618				30	18	11	9	1	1													
603					19	13	10	3		1	1											
570						21	19	7	4	1				1								
94							44	16	8	4	2	3	1									
627								17	9	5	5	5	1									
519									21	23	14	12	3	4	1	1		1				
534										34	33	21	15	12	8	6	3	8			3	
322											25	23	15	15	11	12	6	11	4	2	4	2
200												27	10	17	8	12	3	7	4	3	4	1
196													23	19	19	13	7	14	4	8	8	3
323														34	32	28	17	21	12	7	17	12
221															33	28	24	24	15	11	13	7
300																22	21	22	11	15	19	14
520																	32	37	24	22	19	20
571																		38	22	22	20	12
220																			26	22	18	15
434																				36	29	27
565																					31	26
488																						34

TABLE V

THE NUMBER OF "BETTER" JUDGMENTS IN TABLE IV REDUCED TO PER CENTS OF THE WHOLE NUMBER OF JUDGMENTS

Sample compared	Samples with which comparison is made																				
	580	595	618	603	570	94	627	519	534	322	200	196	323	221	300	520	571	220	434	565	488	177
607		1.4	1.4		1.4																	
580		12.2	4.1	4.1	4.1																	
595			25.7	14.9	4.1	1.4																
618				40.3	24.3	14.9	12.2	1.4	1.4													
603					25.7	17.6	13.5	4.1		1.4	1.4											
570						28.4	25.7	9.5	5.4	1.4												
94							58.1	21.6	10.8	5.4	2.7	4.1	1.4	1.4								
627								23.1	12.2	6.8	6.8	6.8	6.8	1.4								
519									28.4	31.1	18.9	16.2	4.2	1.4								
534										45.9	44.6	28.4	20.3	16.2	10.8	8.1	4.1	10.8			4.1	
322											33.8	31.1	20.3	20.3	14.9	16.2	8.1	14 9	5.4	2.7	5.4	2.7
200												36.5	13.5	23.0	10.8	16.2	4.1	9.5	5.4	4.1	5.4	1.4
196													31.1	25.7	25.7	17.6	9.5	18.9	5.4	10.8	10.8	4.1
323														45.9	43.2	37.8	22.9	28.4	16.2	9.5	23.0	16.2
221															44.6	37.8	32.4	32.4	20.3	14.9	17.6	9.5
300																29.8	28.4	29.8	14.9	20.3	25.7	18.9
520																	43.2	50.0	32.4	29.8	25.7	27.0
571																		51.4	29.8	29.8	27.0	16.2
220																			35.1	29.8	24.3	20.3
434																				48.7	39.2	36.5
565																					41.9	35.1
488																						45.9

Section 8. Results Obtained with the Second Set of Samples

The method employed in selecting the samples that were placed in the second set was given in the preceding section. The set is reproduced below except that samples 519 and 520 were not included in all of the sets judged.

Sample 607. Artificial.

Sketch.

I words four and two came go billa guni sing hay cows and horses he done it good he died it goon I want yes sir yes sir oxes and sheeps he come yes sir camed and goes billum gumun oomunn goodum.

Sample 491. Artificial.

I say never mind. I say he knows all right all right dony puki I say it aint. I got one I got to noo and loo, lov and sov. Dont you care all right all right I say

Sample 58. Artificial.

i from thre ours up in room on books and books care for childer tore a page and rite on them

Sample 501. Artificial.

dere techer:
I like schol not like schol. that man other place like make work tools. Some day you say I rede

Sample 670. Artificial.

I want to say it aint no youse they aint got no right and they aint got no man ban tan pan pan san san sem sen sun sun tun tun

Sample 580. Artificial.

Letter.

Dear Sir: I write to say that it aint a square deal Schools is I say they is I went to a school. red and gree green and

brown aint it hito bit I say he don't know his business not today nor yeaterday and you know it and I want Jennie to get me out.

Sample 97. Artificial.

I write to tell I like the job to work what mr Lambert said as he giv me a dolar harf a day to do things round his plase. I not affrade to work so I say plees giv me the job and I can look after the cow all right or anything no trouble long time I work for mr. panter he knows

Sample 39. Artificial.

The man take little girls in big to the schul when it rains. the wagon with thing to keep warm and light up the rode all the way to Mis Gow.

Sample 595. Artificial.

My Favorite Book.

the book I refer to read is Ichabod Crane, it is an grate book and I like to rede it. Ichabod Crame was a man and a man wrote a book and it is called Ichabod Crane i like it because the man called it ichabod crane when I read it for it is such a great book.

Sample 618. Artificial.

The Advantage of Tyranny.

Advantage evils are things of tyranny and there are many advantage evils. One thing is that when they opress the people they suffer awful I think it is a terrible thing when they say that you can be hanged down or trodden down without mercy and the tyranny does what they want there was tyrans in the revolutionary war and so they throwed off the yok.

Sample 603. Artificial

A Character Sketch.

The man I am describing is a white man and he has nice hair and wears a hat, and his horse is black, I like this man and he has two eyes and his nose is red.

Sample 570. Artificial.

Description of School Room.

Our school room is on the side of the school house and it is a awfully nice room and I like it because it is so nice and all the boys like it, and there is a good many pictures on the wall and there is a clock on the wall. We like this school room and come to school most all the time.

Sample 627. Artificial.

A Scene.

I think the sunlight is very beautiful on the water, and when it shines on the water it is very beautiful, and I love to watch it when it is so beautiful. The colors are so pretty and the noise of the water with the sunshine are so attractive in the sunshine I wonder do other people love to watch the water like I do. I dont know as there is anything as lovely as the water waves in the sunlight of the glorious orb.

Sample 94.

Sulla as a Tyrant.

When Sulla came back from his conquest Marius had put himself consul so sulla with the army he had with him in his conquest siezed the government from Marius and put himself in consul and had a list of his enemys printy and the men whoes names were on this list we beheaded.

Sample 200.

My dear Fred,—

I will tell you of my journey to Delphi Falls, N. Y. There is nice scenery along this route. The prettiest scene is in the glulf which is quite narrow, a small creek flows down it and the road follows along near its banks.

There are woods on either side, these trees look very pretty when they are white with snow.

In summer it is always shady and cool in them and the small fish may be seen darting back and forth in the water.

I hope I will have the pleasure of taking you over the route some time. Yours sincerely,

Sample 322.

A Picture.

I should like to see a picture, illustrating a part of L'allegro. Where the godesses of Mirth and Liberty trip along hand in hand. Two beautiful girls dressed in flowing garments, dancing along a flower-strewn path, through a pretty garden. Their hair flowing down in long curls. Their contenances showing their perfect freedom and happiness. Their arms extended gracefully smelling some sweet flower. In my mind this would make a beautiful picture.

Sample 534.

Fluellen.

The passages given show the following characteristic of Fluellen: his inclination to brag, his professed knowledge of History, his complaining character, his great patriotism, pride of his leader, admired honesty, revengeful, love of fun and punishment of those who deserve it.

Sample 196.

Ichabod Crane.

Ichabod Crane was a schoolmaster in a place called Sleepy Hollow. He was tall and slim with broad shoulders, long arms that dangled far below his coat sleeves. His feet looked as if they might easily have been used for shovels. His nose was long and his entire frame was most loosely hung to-gether.

Sample 221.

Going Down with Victory.

As we road down Lombard Street, we saw flags waving from nearly every window. I surely felt proud that day to be the driver of the gaily decorated coach. Again and again we were cheered as we drove slowly to the postmasters, to await the coming of his majestie's mail. There wasn't one of the gaily bedecked coaches that could have compared with ours, in my estimation. So with waving flags and fluttering hearts we waited for the coming of the mail and the expected tidings of victory.

When at last it did arrive the postmaster began to quickly sort the bundles, we waited anxiously. Immediately upon re-

ceiving our bundles, I lashed the horses and they responded with a jump. Out into the country we drove at reckless speed —everywhere spreading like wildfire the news, "Victory!" The exileration that we all felt was shared with the horses. Up and down grade and over bridges, we drove at breakneck speed and spreading the news at every hamlet with that one cry "Victory!" When at last we were back home again, it was with the hope that we should have another ride some day with "Victory."

Sample 300.

The Preacher of Auburn.

The most popular man of Auburn was the preacher. Although he had a very small salary he was contented. The preacher was kind to everybody. Little children loved him. Old soldiers liked to sit by his fireside and tell stories of the battles, which they had fought in. The beggars who came to his door, although chided for leading such an existence, were always clothed and feed.

The preacher was always willing to go to the homes where there was sickness or death. Here he helped in all things that he could.

In the church he preached with unaffected grace, and all who came to scoff at him remained to worship.

The minister was a contented, simple and kind man, whom the people loved.

Sample 323. Artificial.

Essay on Burns.

As far as I can learn from the Essay on Burns, Mr. Carlyle considers that good poetry must contain the sincerity of the poet. The poem must show the author's good choice of subject and his clearness of sight. In order to have good poetry the poet must be familiar with his subject and his poem will show it.

The characteristics of a great poet, in Mr. Carlyle's opinion were sincerity and choice of subject. A poet must be appreciative of nature and have a responding heart. Carlyle says a true poet does no have to write on subjects which are far away and probably come from the clouds. A truly great poet makes the most of subjects which are familiar to him and close to earth, as Burns did in his poems to the Field Mouse and The Daisey.

Sample 565. Artificial.

October.

I can say without blushing that I think I was born in the best month of the year—October. There is something in a fair October day that makes me feel healthier and happier than any other time of the year. I think there are good reasons for this feeling.

At what time of the year do we have such crisp mornings, such glorious noons and such soft cool twilights? All the sweetness of summer seems to be merging into the chill of winter. There is contentment in the atmosphere. The farmer is rejoicing over his crops; the hunter is seeking satisfaction and pleasure in the forest; the business man is looking forward to the holiday trade; the college man, fresh from his vacation, is at his best; and Nature herself is aflame with color. The whole world seems to be at peace with itself. It is the vacation time for Mother Earth, a lull between the torrid heat of summer and the cold blasts of winter.

Sample 434. Artificial.

A Diary.

I had an early run in the woods before the dew was off the grass. The moss was like velvet and as I ran under the arches of yellow and red leaves I sang for joy, my heart was so bright and the world was so beautiful. I stopped at the end of the walk and saw the sunshine out over the wide "Virginia meadows."

It seemed like going through a dark life or grave into heaven beyond. A very strange and solemn feeling came over me as I stood there, with no sound but the rustle of the pines, no one near me, and the sun so glorious, as for me alone. It seemed as if I had felt God as I never did before, and I prayed in my heart that I might keep that happy sense of nearness all my life.

Sample 177. Artificial.

A Foreigner's Tribute to Joan of Arc.

Joan of Arc, worn out by the suffering that was thrust upon her, nevertheless appeared with a brave mien before the Bishop of Beauvais. She knew, had always known that she must die when her mission was fulfilled and death held no terrors for her. To all the bishop's questions she answered firmly and without hesitation. The bishop failed to confuse her and at

last condemned her to death for heresy, bidding her recant if she would live. She refused and was lead to prison, from there to death.

While the flames were writhing around her she bade the old bishop who stood by her to move away or he would be injured. Her last thought was of others and De Quincy says, that recant was no more in her mind than on her lips. She died as she lived, with a prayer on her lips and listening to the voices that had whispered to her so often.

The heroism of Joan of Arc was wonderful. We do not know what form her great patriotism took or how far it really led her. She spoke of hearing voices and of seeing visions. We only know that she resolved to save her country, knowing though she did so, it would cost her her life. Yet she never hesitated. She was uneducated save for the lessons taught her by nature. Yet she led armies and crowned the dauphin, king of France. She was only a girl, yet she could silence a great bishop by words that came from her heart and from her faith. She was only a woman, yet she could die as bravely as any martyr who had gone before.

Sample 220.

Going Down with Victory.

I sat on the top of a mail-coach in Lombard street impatiently atwaiting the start. 'Twas the night of the victory and we would help spread the news over England.

Up jumps the coachman followed by the guard, an instant's preparation, a touch of the lash and we are off! We are soon past the limits of the city out in open country, galloping, tearing along, a clear road ahead of us for the English Mail stops for nothing.

We dash in at villages, stopping but a moment with the mail, shouting the news of the victory and we are off again. Proud were we and had we not a right to be? The first to carry the great news through the land!

The memory of that ride is ever fresh in my mind and I will ever remember those hours as the most glorious in all my life.

Sample 488. Artificial.

Letter.

Your letter gave me real and heartfelt pleasure, mingled with no small share of astonishment. Mary had previously informed me of your departure for London, and I had not ventured to calculate on any communication from you while surrounded by the splendours and novelties of that great city,

which has been called the mercantile metropolis of Europe. Judging from human nature, I thought that a little country girl, for the first time in a situation so well calculated to excite curiosity and to distract attention, would lose all remembrance for a time at least, of distant and familiar objects, and give herself up entirely to the facination of those scenes which were then presented to her view. Your kind, interesting and most welcome epistle showed me, however, that I had been both mistaken and uncharitable in these suppositions. I was greatly amused at the tone of nonchalance which you assumed while treating of London and its wonders. Did you not feel awed while gazing at St. Paul's and Westminister Abbey? Had you no feelings of intense and ardent interest when in St. James you saw the palace where so many of England's kings have held their courts, and beheld the representations of their persons on the walls? You should not be too much afraid of appearing country-bred; the magnificance of London has drawn exclamations of astonishment from travelled men, inexperienced in the world, its wonders and beauties. Have you yet seen anything of the great personages whom the sitting of Parliment now detains in London—the Duke of Wellington, Sir Robert Peel, Earl Grey, Mr. Stanley, Mr. O'Connell? If I were you, I would not be too anxious to spend my time in reading whilst in town. Make use of your own eyes for the purpose of observation now, and, for a time at least, lay aside the spectacles with which authors would furnish us.

Sample 571.

Venus of Melos.

In looking at this statue we think, not of wisdom, or power, or force, but just beauty. She stands resting the weight of her body on one foot, and advancing the other (left) with knee bent. The posture causes the figure to sway slightly to one side, describing a fine curved line. The lower limbs are draped but the upper part of the body is uncovered. (The unfortunate loss of the statute's arms prevents a positive knowledge of its original attitude.) The eyes are partly closed, having something of a dreamy langour. The nose is perfectly cut, the mouth and chin are moulded in adorable curves. Yet to say that every feature is of faultless perfection is but cold praise. No analysis can convey the sense of her peerless beauty.

More than one hundred of the sets consisting of twenty-seven samples were mailed to individuals whose positions as teachers, authors, and literary workers implied that they were competent

judges of English writing. The following directions accompanied each set: "Please arrange these samples in order of merit as specimens of English composition by young people. After determining your arrangement, pile all the specimens so that the best is on the bottom and the poorest on the top, and securely fasten them. Do not consult with anyone."

When seventy-five replies had been received the results were tabulated in the same manner as in the case of the first set. Tables VI and VII give the same data for this set as are given in Tables III and IV for the first set of judges.

Meanwhile Dr. E. L. Thorndike secured the judgments of forty-one individuals who were especially competent to judge merit in English writing. In order that these results might be used as a check on the others, they are tabulated separately in Tables VIII, IX, X, and XI.

TABLE VI

THE DISTRIBUTIONS OF THE POSITIONS ASSIGNED TO THE SAMPLES IN A SET OF TWENTY-SEVEN BY SEVENTY-FIVE INDIVIDUALS

Positions	Samples																									
	607	491	58	501	670	580	79	39	595	618	603	570	94	627	534	322	200	196	323	221	300	571	220	434	565	488	177
1	38	18	13	4	1																						
2	18	14	22	14	7																						
3	12	13	18	19	7	2	3	1																			
4	2	14	11	22	15	4	4	1	1																		
5	3	6	8	11	23	14	7	2	1	2	1																
6	1	6		3	15	19	18	5	4	1	1																
7	1	3	1	1	4	17	25	13	5	1	3			1													
8					1	10	9	25	13	7	6	2	2														
9		1			1	5	5	11	18	8	15	8	1	2													
10			2			3	1	6	16	19	11	8	7	1		2											
11				1	1	1	2	4	12	13	11	13	9	5	2	1											
12							1	2	4	10	16	12	20	8			2										
13								2	1	8	8	17	16	11	7	3	1	1									
14								1		4	2	7	12	25	6	4	8	4	1			1		1			
15												4	5	10	18	10	10	7	2	2		1	2		1		
16								1		1	1	2	1	7	12	13	7	13	6	6	2	3		1	1		
17								1				1		2	8	14	11	12	5	7	2	5	2	2	2		
18													1	2	5	7	8	19	12	3	5	1		3	3	1	3
19												1			5	5	7	10	8	8	10	5	3	6	4	2	1
20														1	6	4	3	5	9	13	5	8	12	5	4	1	
21															2			1	11	9	16	7	8	6	3	5	5
22																5	2	2	6	8	11	14	6	4	5	7	3
23														1		1	5	1	6	8	5	12	13	9	4	5	6
24															3	3	4		2	5	7	9	13	6	10	5	9
25															1	3	4		7	1	2	4	8	11	9	14	11
26																	3			2	7	3	5	9	15	20	10
27																				3	1	2	3	12	14	15	26

TABLE VII

THE NUMBER OF "BETTER" JUDGMENTS OF SEVENTY-FIVE INDIVIDUALS WITH A SET OF TWENTY-SEVEN SAMPLES

The table should be read as follows: 13 of the seventy-five individuals considered sample 607 better than sample 491, 19 considered sample 491 better than sample 58, 47 considered sample 491 better than sample 58, etc.

Sample compared	Samples with which comparison is made																								
	491	58	501	670	580	79	39	595	618	603	570	94	627	534	322	200	196	323	221	300	571	220	434	565	488	177
607	13	19	10	20	1	1	3	1																		
491		47	18	48	3	2	7	1	1																	
58			10	36	6	3	1	2	1	2																
501				54	16	3	10	5	3	3	1															
670					6	3	7	3	1																	
580						19	31	8	5	6	1		11													
79							62	28	16	15	6	9	3	3	1		2	1								
39								12	6	7	2	2			2		2									
595									25	20	5	10	3													
618										37	17	24	9													
603											19	22	9	3	2		1									
570												42	21	4	4	2	3	1	1	1	1	1	1	1		
94													19	6	5	3	1		1	1		1				
627														15	11	5	12	1	2	1			1			
534															31	27	21	13	13	11	4	10	5	3	6	1
322																34	33	21	13	12	10	14	10	5	6	4
200																	39	12	14	7	4	7	4	1	6	2
196																		24	19	12	12	19	10	6	12	8
323																			33	24	20	26	18	2	15	7
221																				34	24	30	21	14	18	6
300																					26	30	23	8	21	16
571																						44	33	27	20	23
220																							26	18	23	15
434																								24	33	31
565																									47	38
488																										31

TABLE VIII

THE DISTRIBUTIONS OF THE POSITIONS ASSIGNED TO THE SAMPLES IN A SET OF TWENTY-SEVEN AND JUDGED BY FORTY-ONE INDIVIDUALS

Positions	Samples																									
	607	670	58	491	501	580	39	79	595	603	618	94	570	627	534	322	196	200	323	221	300	220	571	434	488	177	565
1	26	10	3	1		1																					
2	9	9	12	7	3	1																					
3	3	9	9	16	5		3																				
4	2	5	8	12	7	4	2																				
5		5	5	3	1	10	3	1						1													
6		3	4	1	12	8	10				3																
7				1	2	10	14	3	6	1	3																
8	1					4	3	16	14	2	2																
9					1	2	4	11	10	1	9	2															
10						1	2	5	11	11	10	6															
11								1	5	8	7	10	5	5													
12								3		7	6	9	12	2	1			1									
13								1		5	1	8	19	4	2	1	1										
14										4		1	5	14	6	2	3	2	2								
15										2		3		6	14	3	3	9	1								
16												1		7	5	6	9	5	3	1			3		1		
17												1		1	4	5	5	7	8	2	5	1	1	1	2		
18														1	6	8	7	7	5	2	3	1	1				
19															3	2	4	4	6	8	6	1	3		1	2	1
20																6	5	2	3	5	5	2	3	2	2	2	4
21																1	2	1	3	7	8	7	4	3	2	1	2
22																4	2		5	4	7	8	1	4	3	2	3
23																1		1	2	1	4	6	3	6	3	7	8
24																1		2		2	2	3	7	11	8	4	2
25																1	1	3	2	3		5	5	4	1	7	10
26																				5	1	2	4	6	8	8	7
27																				1		7	6	4	10	9	3

The author is indebted to Dr. E. L. Thorndike for this table.

TABLE IX

THE NUMBER OF "BETTER" JUDGMENTS OF FORTY-ONE INDIVIDUALS WITH A SET OF TWENTY-SEVEN SAMPLES

For the proper reading of this table see Table VII.

Sample compared	Samples with which comparison is made																								
	491	58	501	670	580	79	39	595	618	603	570	94	627	534	322	200	196	323	221	300	571	220	434	565	488	177
607	8	8	3	12	2		2																			
491		19	10	29	2		3																			
58			7	24	7		2																			
501				31	15		5	1																		
670					4		2																			
580						5	18	5				1														
79							39	20	13	5		3														
39								5	7	2																
595									12	3		5	1													
618										14	3	7	2		1											
603											10	17	6	2												
570												28	10	1												
94													10	2	1		2									
627														11	3	5	6									
534															14	15	11	3	1	1						
322																24	24	20	10	10	9	3	3	3	3	2
200																	19	13	5	5	5	3	2	2	4	2
196																		15	11	8	3	2	2	1	3	2
323																			10	12	4	8	4	4	3	5
221																				23	16	10	13	13	9	8
300																					9	10	4	6	8	5
571																						20	17	18	15	13
220																							19	17	16	11
434																								21	16	18
565																									17	21
488																										19

The author is indebted to Dr. E. L. Thorndike for this table. (Essential data only.)

TABLE X

THE NUMBER OF "BETTER" JUDGMENTS GIVEN IN TABLE XI REDUCED TO PER CENTS OF THE TOTAL NUMBER OF JUDGMENTS

Sample compared	Samples with which comparison is made																								
	491	58	501	670	580	79	39	595	618	603	570	94	627	534	322	200	196	323	221	300	571	220	434	565	488	177
607	19.5	18.5	7.5	29.3	4.9		4.9																			
491		46.3	24.4	70.7	4.9		7.3																			
58			17.1	58.5	17.1		4.9																			
501				75.6	36.6		36.6	2.4																		
670					9.7		4.9																			
580						12.2	43.9	12.2				2.4														
79							92.7	48.8	31.7	12.2		7.3														
39								12.2	17.1	4.9																
595									29.3	7.3		12.2	2.4													
618										34.1	7.3	17.1	4.9													
603											24.4	41.4	14.6	4.9												
570												68.2	24.4	2.4												
94													24.4	4.9	2.4		4.9									
627														26.9	7.3	12.2	14.6									
534															34.1	36.6	26.9	7.3	2.4	2.4						
322																58.5	58.5	48.8	24.4	24.4	21.9	7.3	7.3	7.3	7.3	4.9
200																	46.3	31.7	12.2	12.2	12.2	7.3	4.9	4.9	9.7	4.9
196																		36.6	26.9	19.5	7.3	4.9	4.9	2.4	7.3	4.9
323																			24.4	29.3	9.7	19.5	9.7	9.7	7.3	12.2
221																				56.1	39.0	24.4	31.7	31.7	21.9	19.5
300																					21.9	24.4	9.7	14.6	19.5	12.2
571																						48.8	41.4	43.9	36.6	31.7
220																							46.3	41.4	39.0	26.9
434																								51.2	39.0	43.9
565																									41.4	51.2
488																										46.3

TABLE XI

THE DIFFERENCES BETWEEN THE VARIOUS SAMPLES IN A SET OF TWENTY-SEVEN SAMPLES DERIVED FROM THE PER CENTS OF "BETTER" JUDGMENTS GIVEN IN TABLE X, AND EXPRESSED IN TERMS OF THE DIFFERENCE WHICH EXACTLY SEVENTY-FIVE PER CENT OF THE JUDGES ARE ABLE TO DISTINGUISH (see Section 5)

Sample compared	Samples with which comparison is made																								
	491	58	501	670	580	79	39	595	618	603	570	94	627	534	322	200	196	323	221	300	571	220	434	565	488	177
607	140	99	165	93																						
491		−47	105	−53	260	289	196																			
58			165	7	208	260																				
501				−86	118	260	165	224	260																	
670					208	260	196	260																		
580						99	39	185	224	208																
79							−139	48	118	125	208	174	260	260												
39								147	208	196																
595									64	93	224	165	260													
618										3	112	69	174													
603											99	81	174	260												
570												−22	86	240	240											
94													99	208	224	260										
627														125	156	224	147									
534															34	53	86	140	140	156	240	165	224	260	208	
322																19	22	86	140	147	165	132	165	224	208	240
200																	−7	147	132	196	240	196	240		208	
196																		69	99	147	147	99	165	208	147	185
323																			22	69	93	59	105		125	196
221																				19	69	38	86	132	105	208
300																					59	38	76	185	86	118
571																						−33	22	53	93	76
220																							59	105	76	147
434																								69	22	34
565																									−47	−2
488																										34

A preliminary examination of the results obtained with the set of twenty-seven samples indicated that two other samples should probably have been added. These samples, 519 and 520, were added and twelve individuals judged this set of twenty-nine samples. The distribution of positions assigned the various samples and the number of "better" judgments are given in Tables XII and XIII.

After two hundred and two individuals had judged one or the other of the sets of samples, it was decided to form the scale. Twenty-one of the samples were common to each of the sets. The number of "better" judgments for these samples was obtained by combining the data bearing on these samples from Tables IV, VII, IX, and XIII. The "better" judgments were reduced to per cents and the differences between the various samples determined and given in Tables XV and XVI respectively. The samples composing the scale as given in Section 2 were selected as a convenient series from the set given in Table XVI.

TABLE XII

DISTRIBUTION OF THE POSITIONS ASSIGNED TO EACH OF TWENTY-NINE SAMPLES BY TWELVE INDIVIDUALS

Positions	Samples																											
	607	670	58	491	501	580	39	79	595	603	618	94	570	627	519	534	322	195	300	323	221	300	520	220	571	434	488	177	565
1	8	3	1																										
2	2	4	2	4																									
3		3	4	3	1	1																							
4		1	3	3	4		1																						
5	2	1	2	2	4	1																							
6					2	2	5	2			1																		
7					1	6	3	1	1																				
8						1	2	5	2	1	1																		
9						1	1	2	2	2	1	3																	
10									5	2	3	1	1																
11								1	1	1	2	4	2				1												
12									1	3	1	1	1	4	1														
13								1		3	1	1	1	2	3														
14											2	2	3	2		2					1								
15													1	1	3	4	2	1											
16													1	1	1	2	2	1	2	2									
17														1	3		2	3	2	1									
18													2	1			1	2	4				1	1					
19															1	2		2	1	4							1		1
20																	3			1	4	1		1	1				1
21																		1	2	1	1	5	2						
22																1				2	1	1	1	2	2	1			
23																1		1			3		2	2		1			2
24																		1		1		2			1		2	2	3
25																			1			1	3	1	2	2	1	1	
26																								4	2	2		1	3
27																					2	2	1			3	3		1
28																							1	1	3	3	2	2	
29																							1		1		3	6	1

TABLE XIII

THE NUMBER OF "BETTER" JUDGMENTS OF TWELVE INDIVIDUALS WITH A SET OF TWENTY-NINE SAMPLES

For the proper reading of this table see Table VII.

Sample compared	Samples with which comparison is made																										
	491	58	501	670	580	79	39	595	618	603	570	94	627	519	534	322	200	196	323	221	300	520	571	220	434	565	488	177
607	2	3	2	3																								
491		6	2	8	1																							
58			1	10	1																							
501				10	2																							
670					0	0	0																					
580						3	6																					
79							9																					
39								4	1																			
595									5	4																		
618										6	3	5	3															
603											3	8	2															
570												11	6	4	3	3	1											
94													3	1														
627														5	4	3	1											
519															3	3	1											
534																5	3	2	3	2	1	2	1	1				
322																	3	5	5	2	1	2						
200																		6	3	1	1	1	1		1	1		
196																			3	3	1	2	1	2			1	
323																				3	3							
221																					6	2	2	4	2	2	1	
300																						5	4	5	1	4		1
520																							3	7	2	4	4	1
571																								7	5	7	4	3
220																									4	5	3	1
434																										9	6	4
565																											3	3
488																												5
177																												

TABLE XIV

THE NUMBER OF "BETTER" JUDGMENTS FOR TWENTY-THREE SAMPLES JUDGED BY TWO HUNDRED AND TWO INDIVIDUALS

This table is obtained by combining Tables IV, VII, IX, and XIII.

Sample compared	Samples with which comparison is made																				
	580	595	618	603	570	94	627	519*	534	322	200	196	323	221	300	520*	571	220	434	565	488	177
607	3	2	1																			
580		22	8	9																		
595			61	38	8	17																
618				86	41	47	23															
603					51	60	27															
570						102	56		12													
94							76	48.5	16	10												
627								62	39	22	16	23										
519*									46		31	25										
534										84	78	55	35	28	21							
322											86	85	61	40	34							
200												91	38	37	21							
196													66	52	40	60	23	37				
323														80	71	80	41	55	34	13	35	24
221															96		66	68	51	40	41	21
300																96	60	67	39	33	48	36
520*																			58			
571																		109	77	74	59	51
220																			75	62	60	39
434																				90	85	80
565																					98	89
488																						89

* Samples number 519 and 520 were not included in the set of 27, therefore only 86 of the 202 judges scored them. To keep the number of judges constant, 116 replies to a request printed in *Science* were used.

TABLE XV

THE NUMBER OF "BETTER" JUDGMENTS GIVEN IN TABLE XIV REDUCED TO PER CENTS OF THE TOTAL NUMBER OF JUDGES

Sample compared	Samples with which comparison is made																				
	580	595	618	603	570	94	627	519*	534	322	200	196	323	221	300	520*	571	220	434	565	488	177
607	1.5	1.0	.5																			
580		10.9	4.0	4.5																		
595			30.2	18.8	4.0	8.4																
618				42.6	20.3	23.3	11.4															
603					25.3	29.7	13.4															
570						50.5	27.7		5.9													
94							37.6	24.1	7.9	5.0												
627								30.7	19.3	10.9	7.9	11.4										
519*									22.8		15.4	12.4										
534										41.6	38.6	27.3	17.3	13.9	10.4							
322											42.6	42.1	30.2	19.8	16.8							
200												45.1	18.8	18.3	10.4							
196													39.6	25.7	19.8	29.7	11.4	18.3				
323														39.6	35.2	39.6	20.3	27.2	16.8	6.4	17.3	11.8
221															47.5		32.7	33.7	25.3	19.8	20.3	10.4
300																47.5	29.7	33.2	19.3	16.3	23.8	17.8
520*																			28.7			
571																		54.0	38.1	36.6	29.2	25.3
220																			37.1	30.7	29.7	19.3
434																				44.6	42.1	39.6
565																					48.5	44.1
488																						44.1

* See note with Table XIV.

TABLE XVI

THE DIFFERENCES BETWEEN THE VARIOUS SAMPLES IN A SET OF TWENTY-THREE DERIVED FROM THE PERCENTAGES GIVEN IN TABLE XIV AND EXPRESSED AS MULTIPLES OF THE MEDIAN DEVIATION (Multiplied by 100).

Sample compared	Samples with which comparison is made																				
	580	595	618	603	570	94	627	519*	534	322	200	196	323	221	300	520*	571	220	434	565	488	177
607	—	—	—	—	—																	
580		183	260	252																		
595			77	131	260	204																
618				28	124	109	179															
603					99	79	164															
570						—	88		232													
94							47	105	209	244												
627								75	129	183	209	179										
519*									111		151	171										
534										32	43	90	140	161	186							
322											28	30	77	126	142							
200											—		131	134	186							
196													39	97	126	79	179	134				
323														39	56	39	124	90	142	226	140	176
221															—		66	62	99	126	124	186
300																	79	64	129	145	106	137
520*																			83			
571																			45	51	81	99
220																			49	75	79	129
434																						
565																						
488																						

*See note with Table XIV.

TABLE XVII

THE SAMPLES SELECTED FOR THE SCALE, THE DIFFERENCES BETWEEN ADJACENT SAMPLES AND THE DIFFERENCE BETWEEN EACH SAMPLE AND SAMPLE 580 FOR EACH OF TWO GROUPS OF JUDGES

Sample	Group of 202 judges		Group of 41 judges	
	Differences each from preceding	Differences from sample 580	Differences each from preceding	Differences from sample 580
580		000		000
595	183	183	173	173
618	77	260	81	254
94	109	369	141	395
519	105	474	*	
534	111	585	246†	641
196	90	675	91	732
221	97	772	91	823
571	66	838	41	864
177	99	937	71	935

* This sample is not in the set judged by this group.
† This is the difference between samples 94 and 534.

Section 9. Sufficiency of Judges

No claim is made that the values given in the scale are absolutely accurate. Variation among the judges was very great, and to make a perfect scale would require the services of many more judges than it was possible to secure for this study. The scale is accurate enough to be of very great practical value in measuring the merit of English compositions written in the upper grades of the elementary school and in the high school. The scale will also serve as the basis of future efforts in this direction, and it can be refined and perfected part by part.

The scale was based on the judgment of two hundred and two individuals. Several checks were employed to test the accuracy of the work done by these judges. Table XVI shows the differences between adjacent samples and the absolute values for the samples of the scale. One part of the table shows the results for all the judges and the other part shows the results obtained from the group of judges described on page 36. The differences in the absolute values in the two results are surprisingly small. In no case do the values vary by a complete step in the scale, and the value of the best sample is almost the same in both results. This similarity of values indicates that the values in the scale are reasonably near the values that experts in English composition would assign to the samples.

If the scale were accurate, the values of the samples would remain the same no matter how they were obtained from Table XVI. Thus the difference in value between samples 580 and 618 in the scale is (183 + 77) 260, and in this case the table shows the same difference when sample 580 is compared directly with sample 618. The agreements that exist between the results obtained by these methods are seen in the following tabulation:

Difference	between	samples	580	and	618	in scale is	260,	in Table XVI	260
"	"	"	595	"	94	"	186	"	204
"	"	"	94	"	534	"	226	"	209
"	"	"	534	"	221	"	187	"	161
"	"	"	196	"	571	"	163	"	179
"	"	"	221	"	177	"	165	"	186

Section 10. The Quality of Judges

The question of what individual or group of individuals represents the best judges of English composition is a difficult one to answer. The data at hand do not show any marked differences between the various sets of judges so far as uniformity of judgment is concerned or faithfulness to the proper order of the samples. In order that some definite conclusion might be reached concerning the reliability of judges, the following scale of penalties for placing samples in the wrong order was developed. Where the differences between samples are very great the penalty for misplacing has been made correspondingly large. Thus the difference between samples 627 and 534 is 129. If the judge considers sample 534 better than sample 627, he is guilty of a much greater error than he would be in placing sample 571 before sample 220, where the difference is only 15. In the latter case the penalty was arbitrarily fixed at 1 while in the former it was placed at 6.

TABLE XVIII

The Differences Between Certain Sample Compositions and the Scale of Penalties for Placing Samples in Wrong Order by Individual Judges

Samples compared	Ascertained differences	Penalty for wrong order
94 with 627	47	2
627 " 534	129	6
534 " 322	32	2
322 " 200	43	2
200 " 196	20	1
196 " 323	66	3
323 " 221	39	2
221 " 300	10	1
300 " 220	64	3
220 " 571	15	1
571 " 434	45	2
434 " 565	20	1
565 " 488	5	0
488 " 177	22	1

The following tables show the penalty each judge in each group received and also the median penalty for each group.

TABLE XIX

THE PENALTY FOR WRONG ORDER RECEIVED BY EACH OF THE INDIVIDUALS IN THE SEVERAL GROUPS OF JUDGES

Group of 74				Group of 75				Special Group of 41*	
Judge	Penalty	Judge	Penalty	Judge	Penalty	Judge	Penalty	Judge	Penalty
1	5	39	8	1	7	39	8	1	7
2	5	40	6	2	6	40	7	2	13
3	6	41	8	3	11	41	9	3	10
4	19	42	11	4	5	42	5	4	13
5	7	43	17	5	10	43	10	5	8
6	9	44	5	6	12	44	6	6	8
7	10	45	10	7	9	45	9	7	8
8	8	46	7	8	10	46	11	8	6
9	5	47	9	9	9	47	9	9	10
10	12	48	12	10	12	48	6	10	13
11	6	49	8	11	15	49	9	11	9
12	15	50	8	12	11	50	9	12	16
13	10	51	8	13	8	51	9	13	5
14	10	52	11	14	5	52	4	14	14
15	7	53	11	15	10	53	9	15	9
16	9	54	10	16	10	54	14	16	11
17	9	55	6	17	16	55	3	17	2
18	5	56	7	18	11	56	11	18	13
19	11	57	12	19	14	57	13	19	12
20	8	58	10	20	9	58	7	20	10
21	12	59	15	21	14	59	7	21	12
22	7	60	12	22	8	60	10	22	5
23	9	61	12	23	3	61	11	23	7
24	14	62	12	24	9	62	6	24	9
25	4	63	10	25	16	63	5	25	11
26	4	64	6	26	16	64	8	26	11
27	12	65	7	27	10	65	13	27	7
28	14	66	7	28	13	66	4	28	6
29	9	67	11	29	13	67	8	29	5
30	4	68	11	30	8	68	11	30	13
31	15	69	12	31	9	69	14	31	5
32	12	70	6	32	8	70	15	32	6
33	7	71	8	33	5	71	8	33	2
34	10	72	10	34	7	72	9	34	7
35	11	73	10	35	6	73	13	35	11
36	6	74	11	36	5	74	15	36	14
37	17	75		37	3	75	8	37	6
38	7			38	9			38	

* Data were not calculated for four of the judges in this group.

TABLE XX

DISTRIBUTION OF PENALTIES GIVEN IN TABLE XIX

The table reads as follows: In the group of 41 each of two judges received the minimum penalty, 2; in the group of 75 each of three judges received a penalty of 3; in the group of 74 each of three judges received a penalty of 4, etc.

Penalty	Group of 74	Group of 75	Special group of 41
2			2
3		3	
4	3	2	
5	4	6	4
6	7	5	4
7	9	5	4
8	9	7	3
9	6	12	3
10	10	7	3
11	8	9	4
12	10	3	2
13		5	5
14	2	4	2
15	3	3	
16		3	1
17	2		
18			
19	1		
Median	9.8	9.8	9.6

Table XX shows the distribution of penalties received by each group of judges. It will be recalled that the group of 74 represents the individuals who judged the set composed of 83 samples. These persons were largely graduate students of education. The group of 75 is composed of those who were selected for their general ability as competent judges of English composition. The special group of forty-one is composed of those who may be said to be expert judges of English composition. The table of distribution of penalties shows no important differences in the three groups. The medians for the first two groups are 9.8 in each case while the median for the last is 9.6.

By referring to Table XIX, the individuals whose penalties are at the extremes of the distributions may be found and from data in the possession of the author their identity may be determined. In the special group of 41, judge number 17 received a penalty of 2. This individual is a psychologist, a clear, convincing writer and an editor of large experience. Were his

name mentioned here it would be that of a man who has an international reputation. The largest penalty received by any judge in this group was 16 by judge number 12. All that was said regarding the ability of judge number 17 could be said of this judge, except that his field is philosophy instead of psychology, and he has not had a wide experience in editing. These individuals would be considered equally competent as judges of English composition, yet their judgments show the widest divergence found in this group. Judge number 23 in the group of 75 has a penalty of 3. This individual is a woman who reads a great deal of the best literature. She is not a teacher, and except for occasional papers prepared for a study club, she does not write. Judge number 17 in this same group has a penalty of 16. This individual is the author of several novels and descriptive works. He has had considerable experience as a newspaper reporter and was for some time the editor of a paper. Judge number 6 in the group of 41 is the author of several texts largely used in the English work of the high school.

Section 11. The Use of Mixed Types of Composition

Two methods of selecting samples from which a scale might be derived are possible. The first is to take only compositions that would be classified as narration or as description. The second method is to pay no attention to such distinctions; rather to make the effort to include all the various types. Either method is good. In this study the effort was made to include all types except poetry. It was believed that in this way the scale could be made of more value to the teacher. In the upper elementary grades and in the secondary schools the distinction between narrative, description, and argument is not an important one. Thus the scale will measure the actual products obtained in these grades. The objection may be offered that it is impossible for a person to compare a sample of narration with a sample of description, that we are trying to compare entirely different things. The answer is that people actually did do it. Of the four hundred and fifty people who have judged these samples not more than three have offered any objection on the score that they could not compare the samples.

VITA

Milo Burdette Hillegas was born in the town of St. Johnsville, N. Y., in 1872. He attended the rural and village schools until he entered Colgate academy in 1889. In 1893 he was graduated from this institution and entered the University of Rochester. In 1897 he was graduated with the degree of Ph. B. In 1898 he began teaching in the high school of Ft. Dodge, Iowa, where he remained for two years. He then became vice-principal of the high school at Gouverneur, N. Y., remaining there until 1902, when he became principal of the union schools at Schuylerville, N. Y. In 1904 he became the principal of the union schools in Chatham, N. Y., where he remained until 1908, at which time he entered Teachers College as a candidate for the doctorate. During the second year of his study, he was assistant in the Department of Psychology. In 1910 he was appointed editor-in-chief in the United States Bureau of Education. In June 1911 he resigned this position to accept the position of assistant professor of Elementary Education in Teachers College.

7111

A SCALE FOR THE MEASUREMENT OF QUALITY IN ENGLISH COMPOSITION BY YOUNG PEOPLE

By MILO B. HILLEGAS

Assistant Professor of Elementary Education, Teachers College
Columbia University

Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy, in the Faculty of Philosophy, Columbia University

[Reprinted from TEACHERS COLLEGE RECORD, Vol. 13, No. 4, September, 1912, pp. 331-384.]

PUBLISHED BY
Teachers College, Columbia University
NEW YORK CITY
1912

7111

LB Ap '13

LIBRARY OF CONGRESS
0 021 774 189 0